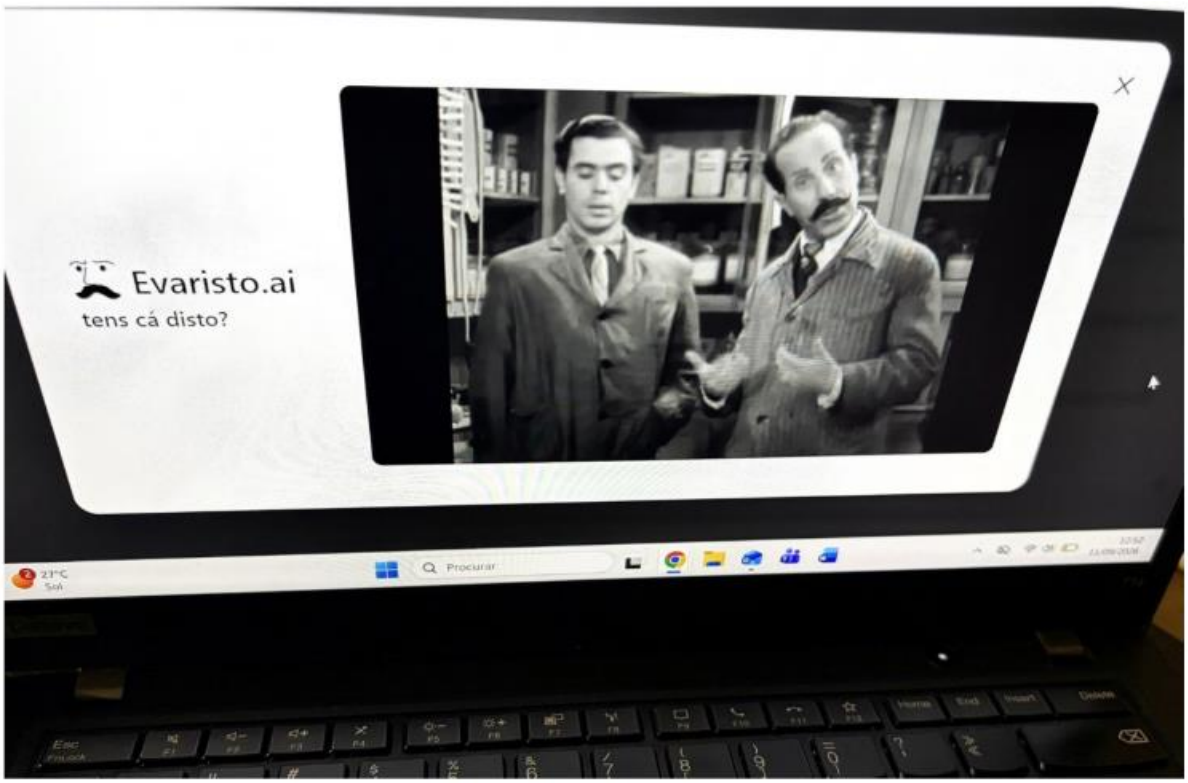
https://www.publico.pt/2026/09/11/enter/noticia/amalia-ja-evaristo-2184232

EXCLUSIVE: ARTIFICIAL INTELLIGENCE

Before Amália, there was Evaristo

In fact, by the time the government unveiled "Portuguese ChatGPT," a team at FCUL had already developed several artificial intelligence models and chatbots for the Portuguese language.

Laura Ferreira
September 11, 2026, 16:13



O nome do *chatbot* foi inspirado no clássico do cinema português, *O Pátio das Cantigas* DR

Before the government introduced Amália as the first open-source language model developed in European Portuguese, there was already an artificial intelligence (AI) ecosystem in Portugal designed for the Portuguese language.

NLX—the Natural Language and Speech Group in the Department of Computer Science at the Faculty of Sciences of the University of Lisbon (FCUL)—has been working for 25 years on technology and AI tailored to the Portuguese language, and three years ago it officially began launching language-based systems.

António Branco, a professor at the University of Lisbon and coordinator of NLX, acknowledges the importance of Amália to the open AI landscape in Portugal, but says it is nothing new. “I think Amália could have made a better contribution by offering something that complemented what was already out there,” he tells PÚBLICO.



António Branco, professor da Universidade de Lisboa e coordenador do NLX DR

The history of “Portuguese ChatGPT,” developed by the Telecommunications Institute, the Instituto Superior Técnico of the University of Lisbon, and the New University of Lisbon, has been marked by controversy from its announcement in late 2024 through its launch in July 2026.

There was discontent surrounding the selection process for the research centers that would carry out the project, questions about the data used to train the model, and concerns about the budget, which rose from the initially projected 5.5 million euros to about seven million.

In the weeks following the launch, comments and criticism on social media multiplied regarding the model’s absurd answers to seemingly simple questions. In one instance, Amália identified Roberto Martínez as the coach of Belgium; in another, she claimed that the word "morango" (strawberry) contained the letter R three times.

“The backlash on social media against Amália—and, by extension, against generative AI—was the last thing we needed,” says António Branco.

When asked whether the investment of about seven million euros in Amália was

justified, António Branco replied: “What I can say is that what we did—eight open models, three chatbots, and several datasets—cost about one-third of that amount.”

Albertinas, Gervásios, Serafins and Evaristos

It was in 2023 that NLX launched the first AI model designed for the Portuguese language: Albertina. The LLM (large language model) belongs to the encoder family—tools specialized in reading, understanding, and classifying text. In a movie review app, for example, when presented with the sentence “I really liked this movie,” Albertina can interpret the meaning and identify that it is a positive review.

Next came Gervásio, from the family of decoders. Unlike the earlier models, these models do not merely interpret the information they receive: they can generate text from it. These models form the basis of many chatbots.

In 2024, Serafim, a vectorizer, was introduced. This type of model is capable of measuring the semantic proximity between words, phrases, or documents. It is a useful technology, for example, in a search engine, which can thus find the semantically closest content and provide the answer to a specific question.

“The next step was to create an interface that would allow us to interact with these models,” recalls António Branco. And that’s when, in June 2025, Evaristo.ai was launched.

Qual é a diferença entre um modelo e um *chatbot*?

Um modelo de linguagem é a tecnologia que aprende padrões a partir de grandes quantidades de dados e consegue, por exemplo, gerar ou compreender texto.

Um *chatbot* é uma aplicação que usa um ou mais modelos de IA para interagir com as pessoas através de uma conversa. É a "interface" que permite fazer perguntas e receber respostas.

O Amália é um modelo de linguagem; O Evaristo é um *chatbot*.

The name was inspired by the Portuguese film classic, *O Pátio das Cantigas*. In the film, Evaristo is the owner of the neighborhood grocery store, a character who became associated with the phrase “Hey, Evaristo, do you have this here?” The project revives this reference for the age of artificial intelligence with the variation: “Hey, Evaristo.ai, do you have an answer for this?”

“I’m sure Evaristo doesn’t know who Roberto Martinez is either, but he’s capable of looking it up and giving the right answer,” says the teacher, referring to

Amália's answers.

Both Amália and the models used by Evaristo were trained using information collected up to a certain date—neither, therefore, has any knowledge of events that occurred after that date. However, Evaristo gets around this problem through what is known as "agentivity": the chatbot's ability to decide when it needs to use external tools, such as the Internet, to complete a response.

"When we developed Evaristo, we realized we had to make it work in line with users' expectations," explains António Branco. "And people's expectations are already shaped by their experiences with other chatbots."

The possibility of Amália beginning to search for information on the web is, for João Magalhães, a researcher at Universidade Nova de Lisboa and the project coordinator, a technically demanding development. "The 'evolution' toward using web search is actually a very challenging application of LLM because the questions are completely open-ended," he said. Still, he admits that applications with this capability may emerge: "Although it is very complex and costly from an engineering and maintenance standpoint, someone will eventually do it."

Amália is an open, general-purpose language model designed to be integrated into and adapted for different applications. It was not launched as a chatbot, like Evaristo. The interfaces that emerged later—which allowed users to converse with the model—were developed by third parties.



The Path to Sovereign AI

The difference between the two projects lies not only in how they can be used, but also in the scope and origin of the models on which they are based.

The most powerful Gervásio model has 70 billion parameters, and the Amália model has 9 billion. These are impressive numbers, but modest when compared to the largest models developed by the major tech companies.

To be trained, LLMs require colossal amounts of text, computing power, and investments that only a few of the world's largest technology companies can afford. "With the open-source data currently available for the Portuguese language,

we would never have been able to achieve the results we have today," António Branco emphasizes.

One way to overcome this obstacle is to start with a pre-trained model and adapt it to a specific language or task. Researchers at FCUL began with Meta's Llama model and continued training it with data in European Portuguese, in order to teach the system the grammatical peculiarities, vocabulary, and expressions specific to the national variant.

Amália is based on EuroLLM, a family of LLMs developed in Europe for various European languages. "We believe that using models from outside Europe, developed behind closed doors, does not truly contribute to the AI ecosystem or national sovereignty," says João Magalhães. "The EuroLLM that powers Amália is completely open-source, and great care was taken to release a model with a truly open license that would protect applications using Amália."

For António Branco, sovereignty does not depend on that. "For an organization to have a sovereign AI service, the software does not have to be developed by that organization; nor in that organization's country; nor on that organization's continent," he argues. For the professor at the University of Lisbon, the ability to take a large, pre-trained model and specialize it may be precisely the way to achieve sovereignty. "Having an LLM specialized in European Portuguese based on a pre-trained base model is not a disadvantage. It is an advantage and even a necessity if one wants to have an LLM with competitive performance."

Openness is what brings the two approaches together, despite their differences. Both Amália and the models developed by NLX are available on Hugging Face, one of the world's largest repositories of artificial intelligence models, where they can be downloaded, studied, and reused.

"After we released Albertina, we received an email saying that MediAlbertina, a medical app, had been created," recalls Professor António Branco. According to the researcher, the NLX models have been downloaded "at least several hundred thousand times" since the launch of the first LLM.



Membros do NLX trabalharam no desenvolvimento dos modelos de linguagem DR

For the researcher, it is precisely this possibility of reuse that justifies making the models openly available. “Building the models is the most costly part of generative AI. We were fortunate to receive funding to develop them. Now we’re making them available to anyone who wants to reuse them.”

NLX’s AI tools were developed as part of a project funded by the PRR, which concluded at the end of June. Portulan Clarin, a research infrastructure that is part of a broader European network created to support researchers in the field of language sciences, is now responsible for the maintenance and operation of the systems.

Models that can run in a basement

“There’s no doubt that ChatGPT or Gemini are exponentially superior to Evaristo,” admits António Branco. But competing with big tech was never the project’s mission.

Instead of trying to build an ever-larger and more powerful model, the NLX team sought to answer another question: Is it possible to develop useful artificial intelligence services using smaller, specialized models that are affordable enough for universities, companies, or public agencies to use?

> **“O Evaristo tem o propósito de ser uma demonstração empírica de que a IA soberana é viável tecnologicamente e comportável economicamente”, defende o investigador.**

“The data centers of big tech companies cost billions because they have to serve millions of users at the same time. A chatbot like Evaristo receives, at most, a few dozen or a few hundred requests per day,” the professor explains. “These models can run in a basement,” says António Branco—installed on the servers of a university, a company, or a government agency, without relying on massive infrastructure.

That is why, he argues, small, specialized models may be more useful than a single system that tries to do everything. Following this logic, the FCUL team began developing systems for different purposes. Today, there are already three versions of the Evaristo chatbot, each designed for a specific context.

The first to be launched was Evaristo.ai – Generalista, designed for general use and with access to extensions such as Internet search. In November 2025, Evaristo.ai – Public Services was launched, a chatbot specializing in Portuguese public administration services, based on information available on the gov.pt website. Then, in May 2026, Evaristo.ai – Plain Language was introduced, created to help rephrase texts in simpler, more accessible language.

The team wanted to take the work it had developed over the years to the next level. To that end, it created the spin-off Ousar.ai to transform its research into AI solutions for organizations.

The company develops agents that can operate directly on the hardware of the organizations themselves, including in sectors such as finance, tourism, and cybersecurity. Evaristo and Gervásio are some of the models used for demonstration purposes. The idea is for institutions to have their own AI tools, in Portuguese, tailored to their needs, without relinquishing control over the technology or their data.